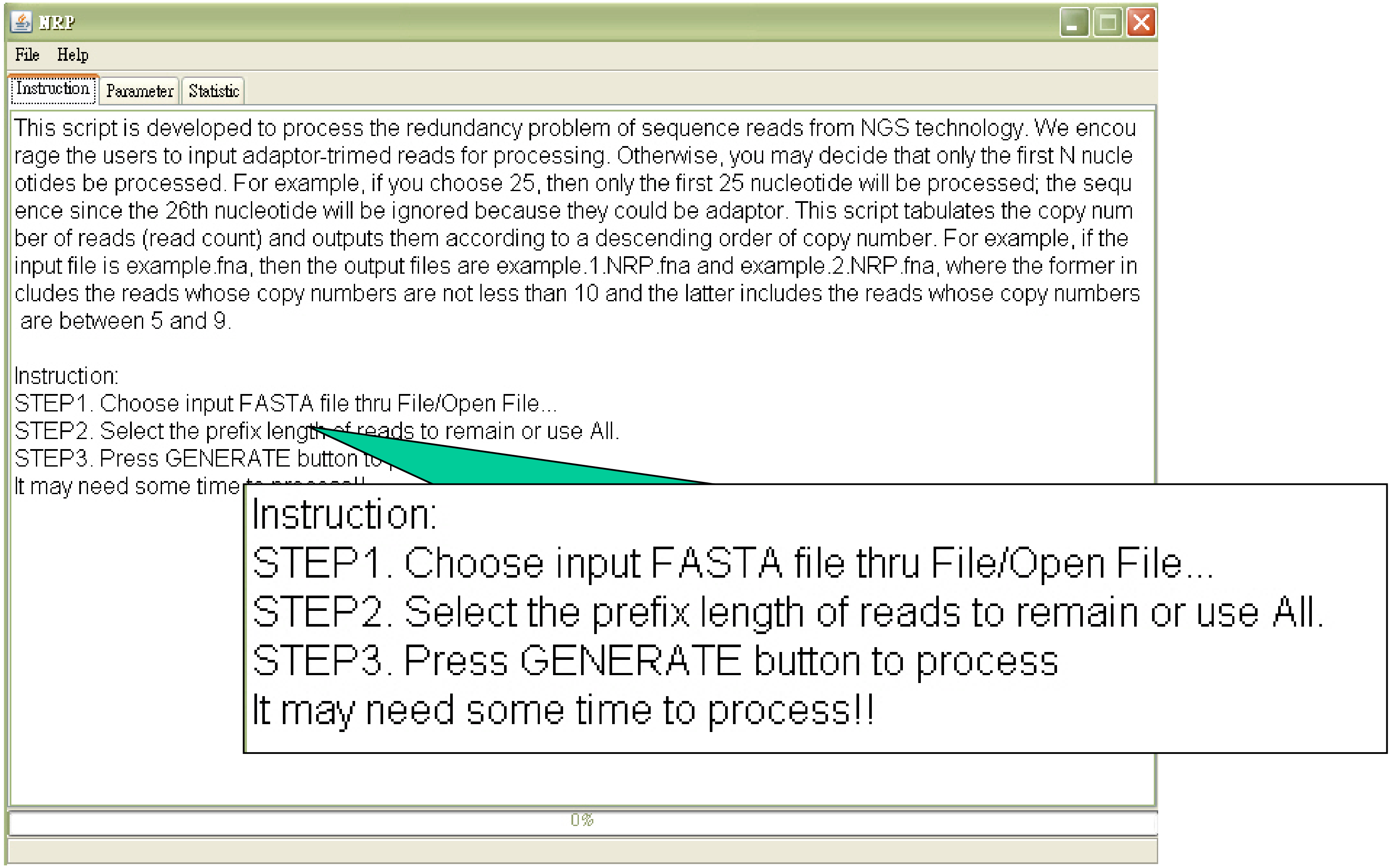
NRP instruction

double click

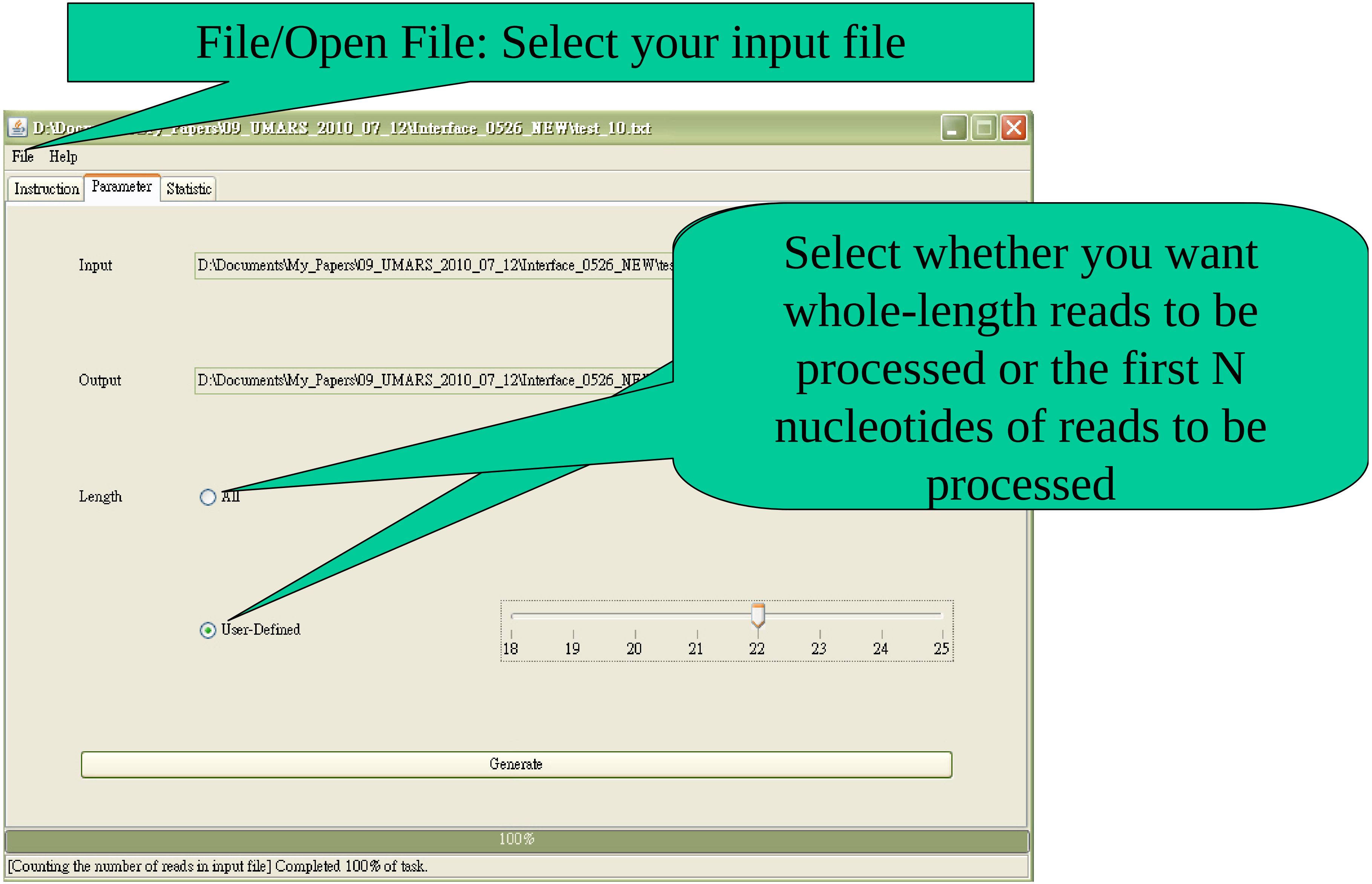


This script is developed to process the redundancy problem of sequence reads from NGS technology. We encourage the users to input adaptor-trimed reads for processing. Otherwise, you may decide that only the first N nucleotides be processed. For example, if you choose 25, then only the first 25 nucleotide will be processed; the sequence since the 26th nucleotide will be ignored because they could be adaptor. This script tabulates the copy number of reads (read count) and outputs them according to a descending order of copy number. For example, if the input file is example.fna, then the output files are example.1.NRP.fna and example.2.NRP.fna, where the former includes the reads whose copy numbers are not less than 10 and the latter includes the reads whose copy numbers are between 5 and 9.

Instruction:
STEP1. Choose input FASTA file thru File/Open File...
STEP2. Select the prefix length of reads to remain or use All.
STEP3. Press GENERATE button to
It may need some time

Instruction:
STEP1. Choose input FASTA file thru File/Open File...
STEP2. Select the prefix length of reads to remain or use All.
STEP3. Press GENERATE button to process
It may need some time to process!!

NRP instruction



File/Open File: Select your input file

Select whether you want whole-length reads to be processed or the first N nucleotides of reads to be processed

Input D:\Documents\My_Papers\09_UMARS_2010_07_12\Interface_0526_NEW\tes

Output D:\Documents\My_Papers\09_UMARS_2010_07_12\Interface_0526_NE

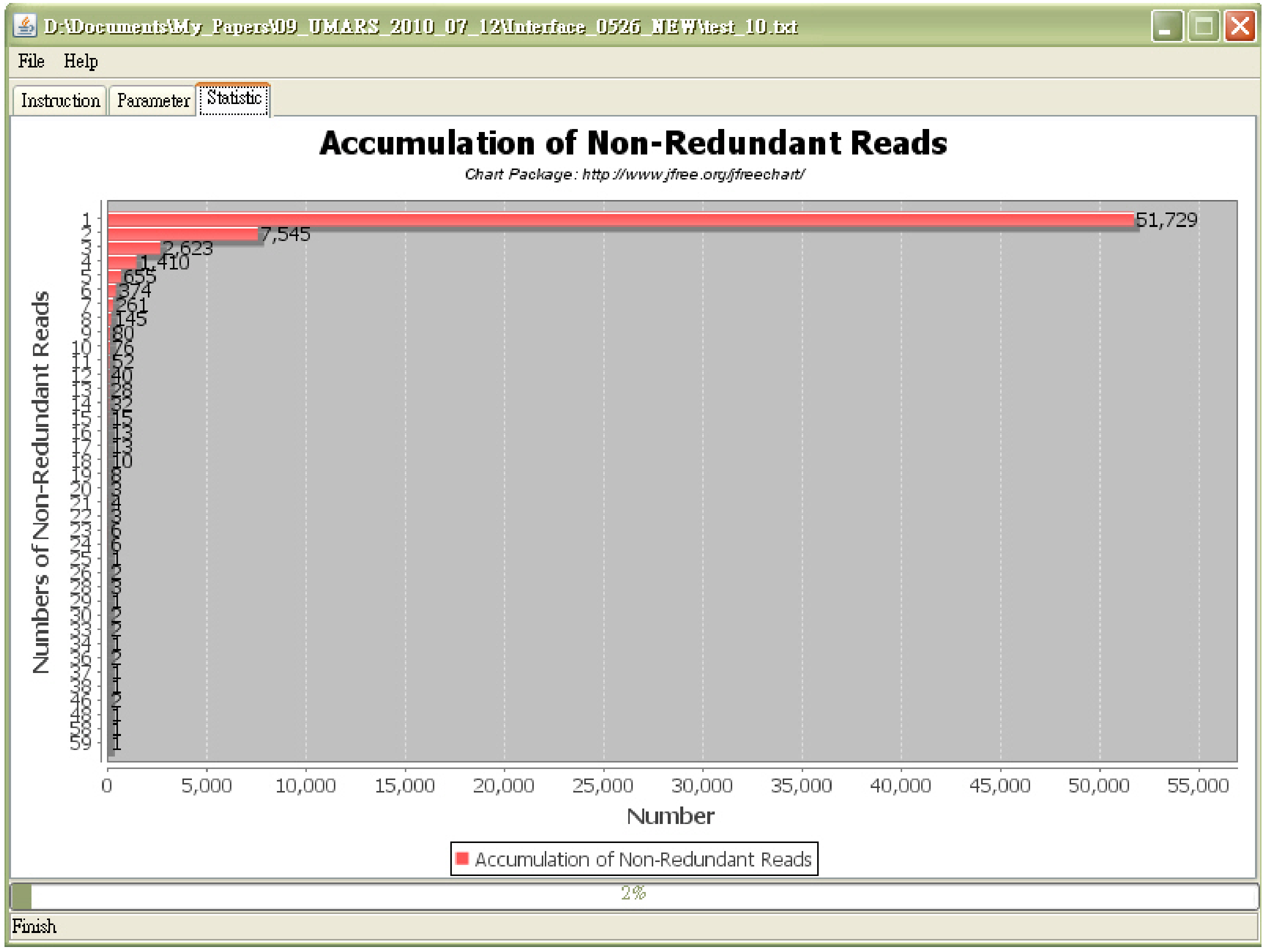
Length ○ All

● User-Defined

100%

[Counting the number of reads in input file] Completed 100% of task.

NRP instruction



Simple statistics of alignment result.

NRP instruction

```
>ILLUMINA:2:1:1024:3541#0/1
AGAAATTTTCTATCTTTATTCAACTTCCATATACTTAATG
>ILLUMINA:2:1:1024:20089#0/1
AGGGATAGATAAAGATATAATCATAAGAATTATATTTTAA
>ILLUMINA:2:1:1025:2375#0/1
CTCATTGAATGGAAATGAAAGGGGTCATCATCTAATGGAA
>ILLUMINA:2:1:1025:9576#0/1
AGACGTATGCAGCTGGTAAATGATGCACACCTAACTGGAT
>ILLUMINA:2:1:1025:12667#0/1
ATGACCATGGCAGCATGGTAATAAAGTGTTATTTACAGAA
>ILLUMINA:2:1:1025:5856#0/1
GGAACAATATATCCCCCTCAAACCCAGTGAGGAGAGAAAC
>ILLUMINA:2:1:1025:18260#0/1
GATGCAACCAGAATTTGGACTAAAGGTACAGAGGGGTGGC
>ILLUMINA:2:1:1026:14066#0/1
ATTCAAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAG
>ILLUMINA:2:1:1026:13797#0/1
GCAGTCTGCAATTTGTATGAATTACTGCTTCCAACGAAAT
```

Original input file

```
>NR1 63
CGCGACCTCAGA[illegible]
>NR2 63
GTCCAGTTTTCCCAGGAATCCC
>NR3 56
GCATTGGTGGTTCAGTGGTAGA
>NR4 54
CGCCTTGGCCGTACAGCAGGGG
>NR5 51
CGCGACCTCAGATCAGACGTGA
>NR6 51
TCCCTGAGACCCTAACTTGTGA
>NR7 46
AGCAGCATTGTACAGGGCTATG
>NR8 45
TGGCTCAGTTCAGCAGGAACAG
>NR9 44
ACGCGACCTCAGATCAGACGTG
```

Output alignment file